Das Umweltforschungs-Informationssystem UFIS - aktueller Stand der Arbeiten im Informationssystem über ökologische Modelle und Daten

C. Herderich, R. Lenz, M. Knorrenschild, O. Springstubbe[1]

1. Zielsetzung und Kurzfassung:

UFIS wird Informationen zur ökologischen Forschung, insbesondere über Modelle und dazugehörige Daten überwiegend aus Projekten des BMFT bereitstellen (sog. Meta-Informationen). Diese Information schließt eine standardisierte Dokumentation von Modellen und solchen Datensätzen ein, die für die ökologische Modellierung von Bedeutung sind und international anerkannten "Protokollen" entsprechen.
Diese Meta-Information wird über ein Nutzer-Führungssystem automatisiert zugänglich sein. Ergänzend werden Analysen, Klassifizierungen und Bewertungen von Modellen durchgeführt, deren Ergebnisse die Dokumentations-Richtlinien optimieren helfen und für den Nutzer eine Art Qualität des Modells erkennbar machen.

Nutzergruppen werden sein:
- Auftraggeber und Projektträger der ökologischen Forschung
- Wissenschaftler, die Modelle entwickeln, nutzen und verbinden wollen, um komplexe Ursache-Wirkungs-Beziehungen in unserer Umwelt (zentrale Aggregationsebene: Ökosystem) simulieren, prognostizieren und bewerten zu können.

Das aktuelle Arbeitsprogramm ist die Inventur und Beschreibung von Forschungsprojekten, die Auswahl von Projekten, die für UFIS Nutzer relevant sind, eine Umfrage zum Modelleinsatz an Ökosystemforschungszentren u.ä. Einrichtungen, sowie die Erstellung eines Dokumentationskonzepts für Modelle und Daten (Klassifizierung, Schlagworte).
Nach bisherigem Kenntnisstand enthält die Datenbank DAVOR des BMFT ca. 900 Projekte für 8 Bereiche der ökologischen Forschung (BMFT Referat 522).

---

[1] GSF - Forschungszentrum für Umwelt und Gesundheit, Projekt Umweltgefährdungspotentiale von Chemikalien, Neuherberg, 85764 Oberschleißheim.
Das diesem Beitrag zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministers für Forschung und Technologie unter dem Förderkennzeichen 0339349A gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

Schlagwort-Abfragen ergaben letztlich 16 Vorhaben, die sich explizit mit Modellsimulationen beschäftigen bzw. beschäftigt haben. In einer Testphase wurden diese Modelle, soweit bereits verfügbar, mit dem vorläufigen Dokumentationsschema erfaßt, und schließlich wurde in gemeinsamer Bearbeitung mit dem Projektzentrum Ökosystemforschung in Kiel und dem Projekt ECOBAS in Kassel eine Modelldokumentation fertiggestellt.
Parallel dazu erfolgte eine Umfrage zum aktuellen und angestrebten Einsatz von Simulationsmodellen an den Ökosystemforschungszentren u.ä. Einrichtungen. Test und Umfrage werden derzeit weiter ausgewertet. Bislang wurden 106 Modelle in eine Kurzbeschreibung aufgenommen.
Für die Erfassung der standardisierten Dokumentation von Modellen wird derzeit ein PC-Programm entwickelt, das den Modellentwickler in die Lage versetzen soll, sein Modell entsprechend zu dokumentieren. Im 2. Quartal 1994 wird ein Datenbankprototyp zur Abfrage der Modelldokumentation verfügbar und über Telnet zu erreichen sein.
Zur Erleichterung der Modelldokumentation bzw., um eine eindeutige Benennung (kontrolliertes Vokabular) der im Modell simulierten Systeme zu gewährleisten, gibt UFIS eine hierarchische Nomenklatur ökologischer Systeme vor. Die Modelldokumentation kann derzeit über ftp eingesehen werden. Für die Metadatenbank zu Daten liegen erste Entwürfe vor.

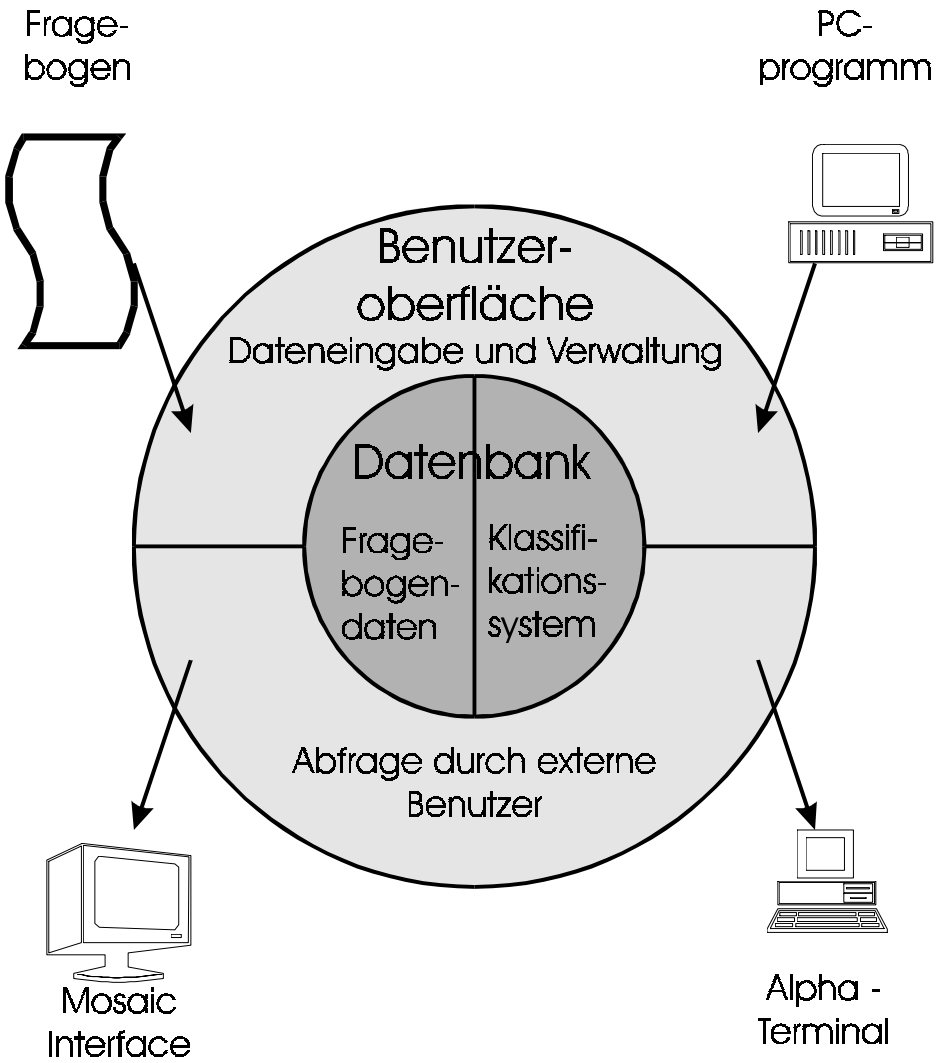
2. Stand der Entwicklung des "Metainformationssystems Modelle"

Ein Schwerpunkt des Projekts UFIS (Lenz u.a. 1993) ist die Entwicklung eines Metainformationssystems über Modelle. Mit Hilfe eines Fragebogens[2] (Benz u.a. 1993) werden standardisierte Modellbeschreibungen erhoben. Das Ziel des Informationssystems ist die Speicherung und Verknüpfung aller in diesem Fragebogen erfaßten Informationen.
Generell kann die Softwareentwicklung in die Phasen Anforderungsspezifikation, Systementwurf, Systemimplementierung und schließlich Systemwartung unterteilt werden. Das Informationssystem selbst besteht aus der eigentlichen Datenbank und den Benutzeroberflächen zur Interaktion mit der Datenbank. Die Datenbank wird weiter gegliedert in den Teil zur Speicherung der Daten, die durch den Fragebogen erhoben werden und in das von UFIS entwickelte Klassifikationssystem (Springstubbe 1993, Lenz 1993). Für die Benutzung der Datenbank ist eine Oberfläche zur Dateneingabe und -verwaltung der papiergebundenen Fragebögen sowie eine Oberfläche zur Datenabfrage für "externe" Benutzer geplant, die auf alphanumerischen Terminals basiert. Damit ist für die Abfrage der Datenbank ein Zugriff

---

[2] Ein Programm, das den papiergebundenen Fragebogen durch eine äquivalente PC-basierte Version ersetzt, ist in Vorbereitung.

über eine Telnetverbindung möglich. Zusätzlich ist geplant, die Informationen der Datenbank im World Wide Web als Hypertextdokument im HTML - Format (Weichselgartner 1993) anzubieten und den Fragebogen als PC-File in die Datenbank einspielen zu können.
Grobstruktur des Metainformationssystems Modelle:



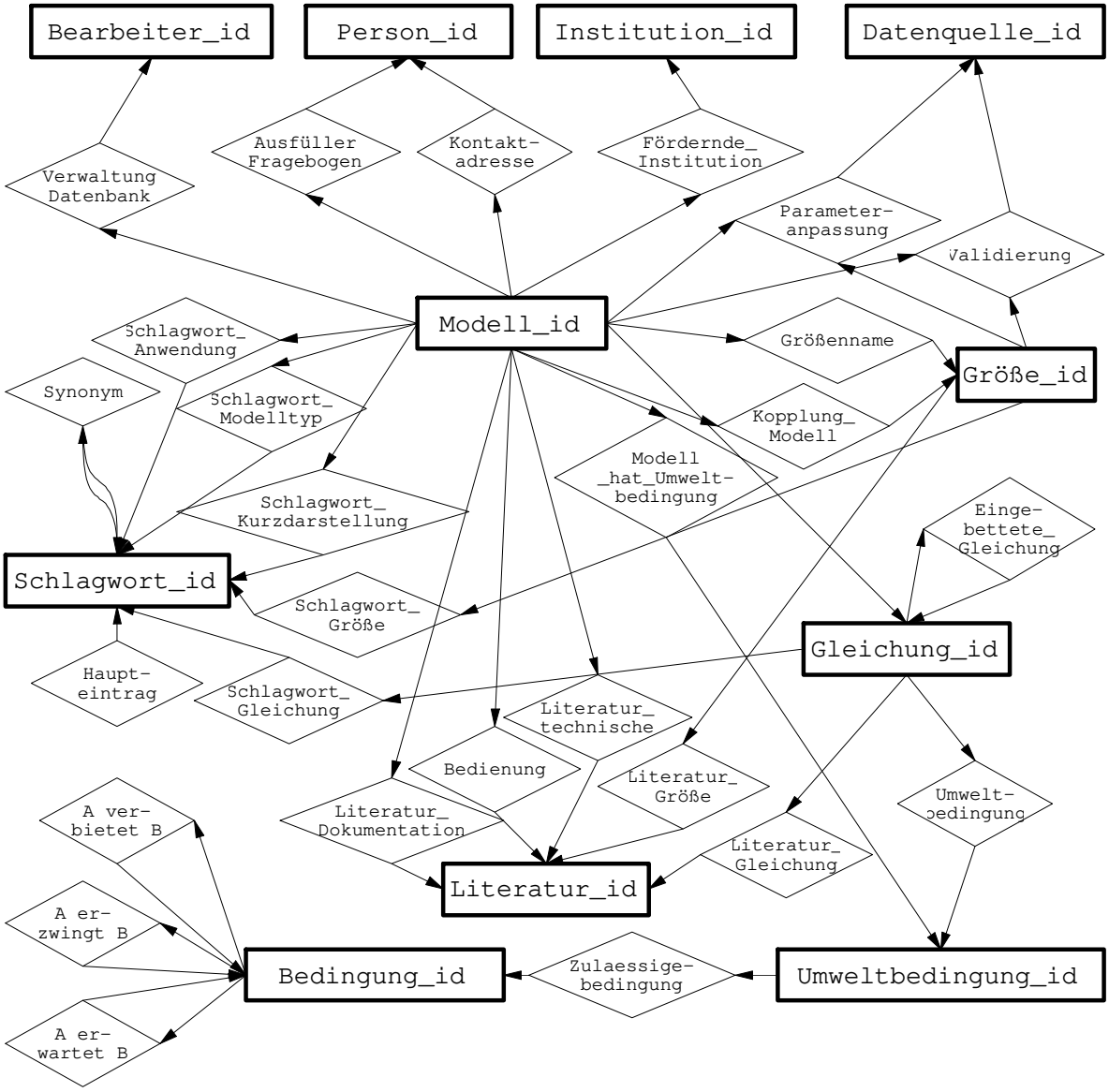
2.1. Anforderungsspezifikation

Die grundlegende Anforderungsspezifikation für alle Systemteile wird durch den UFIS-Fragebogen vorgegeben, der beschreibt, welche Informationen in welcher Form gespeichert werden müssen. Dadurch ist auch implizit festgelegt, wie die

Masken für die Dateneingabe aussehen müssen und nach welchen Informationen bei der Abfrage der Daten gesucht werden kann. Für das Klassifikationssystem wird die Anforderungsspezifikation durch die Arbeit von Springstubbe (1993) und darauf aufbauende Arbeiten ergänzt. Damit ist die Definition der Anforderungsspezifikation im Wesentlichen abgeschlossen. Nach einem noch für 1994 geplanten Probebetrieb werden sich sicher noch Änderungen in der Anforderung an die Datenbank ergeben, die später berücksichtigt werden müssen.

2.2. Systementwurf

Die wesentliche bisher geleistete Arbeit war die Erstellung des Systementwurfs für die Speicherung der im Fragebogen erhobenen Daten. Durch die Verwendung des Datenbanksystems Ingres war vorgegeben, daß das Ergebnis des Systementwurfs ein relationales Datenmodell (Tabellen) sein mußte. Als Vorstufe für ein genaues relationales Modell wurde ein gröberes objektorientiertes Modell entworfen. Als Resultat dieses Arbeitschrittes wurde ein Dokument erstellt, das das relationale Modell mit allen Tabellen, allen Tabellenspalten und ihrer Bedeutung enthält. Zusätzlich wird in diesem Dokument beschrieben, wie sich diese Tabellen zu Objekten zusammensetzen und wie diese Objekte zusammenhängen. Außerdem war noch zu beachten, daß Ingres für einzelne Tabellen und für jede Verknüpfung von Tabellen (das Kreuzprodukt bei Joins) eine maximale Tabellenbreite von ca. 2.000 Zeichen hat. Dies bedeutete eine wesentliche Einschränkung im Systementwurf.
Konkret wurden zunächst größere Informationseinheiten (Objekte) identifiziert. Dann wurde untersucht, in welchem logischen Zusammenhang diese Objekte stehen (Relationen zwischen diesen Objekten). Schließlich wurden die Objekte so lange verfeinert, bis alle Informationen des Fragebogens erfaßt werden konnten. In der folgenden Graphik sind die verwendeten Objekte und ihre Beziehung untereinander dargestellt.
Objekte sind in dieser Graphik immer als <Objektname>_id bezeichnet und als Rechtecke ausgeführt. Relationen sind als Rauten dargestellt. Beim Objekt Modell handelt es sich um die Teile der Beschreibung eines Modells, die nur einmal vorkommen, wie etwa eine Kurzbeschreibung des Modells. Die zugehörigen Objekte Gleichung und Größe beschreiben die im Modell verwendeten Gleichungen und Größen und sind typischerweise in einer Modellbeschreibung mehrfach vorhanden. Die in den Modellen verwendete Literatur, die das Modell charakterisierenden Schlagworte, die das Modell fördernde Institutionen und Personen werden für alle Modelle global gespeichert. Die Objekte Bedingungen und Umweltbedingungen folgen aus dem UFIS-Klassifikationssystem. Verwaltungstechnischer Natur ist das Objekt Bearbeiter, das den Bearbeiter des Datenbankeintrags beschreibt.



Aus der Anforderungsspezifikation läßt sich leicht ersehen, daß eine Aufbereitung der Informationsmenge, die im Fragebogen erhoben wird, in der Art eines Hypertext-Dokuments besonders geeignet ist. Dies bedeutet, die im Fragebogen erhobenen Informationen werden zu einzelnen leicht überschaubaren Informationsblöcken zusammengefaßt, die über sogenannte Links inhaltlich zusammenhängen. Ein Programm, das aus der Datenbank entsprechende Hypertextdokumente ableitet, ist im Entwurf.

2.3. Systemimplementierung

Die Implementierung erfolgt auf einem Convex Rechner mit Unix Betriebssystem unter Ingres Version 6.3. Der bisher vorliegende Systementwurf wurde bislang in Anweisungen für die Datenbank (Ingres - SQL-Anweisungen) umgesetzt und in die Datenbank eingespielt. Die so entstandene Datenbank wird zur Zeit als Arbeits- und Testumgebung bei der Oberflächenentwicklung genutzt. Die alphanumerische Oberfläche wird mit dem INGRES Tool "Application By Form" (ABF) mit der Datenbanksprache INGRES4GL durchgeführt.

2.4. Weitere Arbeitsplanung

Derzeit wird der Systementwurf und die Systemimplementierung für das Klassifikationsschema abgeschlossen. Außerdem ist geplant, die Eingabeoberfläche fertigzustellen und eine erste Testversion für die Datenabfrage zu implementieren. Insgesamt wäre damit ein erster Probebetrieb möglich. Da erfahrungsgemäß während eines Probebetriebs Probleme auftreten, die bis in den Bereich der Anforderungsspezifikation hinein reichen, ist für Ende 1994 eine Redesignphase geplant, nach der ein Dauerbetrieb der "Metadatenbank Modelle" möglich sein sollte.
Im weiteren Verlauf des Projekts wird untersucht, wie Daten beschrieben sein müssen, um eine automatische Kopplung von Daten und Modellen möglich zu machen. Dabei wird dem Aspekt der räumlichen und zeitlichen Auflösung von Daten besondere Beachtung geschenkt (Weisemann 1994). In erster Linie zielen diese Arbeiten auf die automatische Kopplung von Modellen untereinander, da mit Daten zunächst solche Daten gemeint sind, die von anderen Modellen erzeugt werden. Die Situation für die Kopplung von Modellen mit Meßdaten ist aus der Sicht von Modellen jedoch nicht grundsätzlich anders, so daß diese Arbeit ein weiterer Schritt in Richtung auf eine "Metadatenbank Daten" darstellt.
Bislang wurde hierzu eine vierstufige Metainformationsbeschreibung für die Eingabe- und Ausgabegrößen eines Modells entworfen. Die erste Ebene ist eine symbolische Beschreibung der Größe mit Größenname, Klassifikation und Einheiten. Die zweite Ebene ist eine zusammenfassende Beschreibung der räumlichen und zeitlichen Struktur einer Größe. Die dritte Ebene beschreibt die Struktur, d.h. die Auflösung in den Raumrichtungen x, y und z sowie die zeitliche Auflösung (t) genau. Auf der vierten Ebene wird schließlich für jedes Tupel (x,y,z,t) der dritten Ebene ein konkretes Werte - Intervall angegeben, in dem die verwendeten Werte liegen müssen. Auf der Basis dieser vierstufigen Beschreibung wurden Prozeduren entwickelt, die schrittweise feststellen, ob der Output eines Modells als Input für ein weiteres Modell geeignet ist oder nicht. Die Implementierung dieser Konzepte ist für 1994 geplant. Grundsätzlich ist festzuhalten, daß solche Prozeduren immer nur formal prüfen können, ob die Kopplung zweier Modelle möglich ist. Die

semantische Überprüfung muß der Modellentwickler oder der Modellbenutzer selbst durchführen.

Literatur

Benz, J./Knorrenschild, M. (1993): Anforderungen an die Dokumentation mathematischer Beschreibungen ökologischer Prozesse, in: Keller, H. B. u.a. (Hrsg.), Bericht vom 2. Treffen des AK 'Werkzeuge für Simulation und Modellbildung in Umweltanwendungen'. Bericht des Kernforschungszentrums Karlsruhe KfK Nr. 5159. Karlsruhe, S. 9-16

Lenz, Roman (1993): Systemökologische Anforderungen an ein Umweltforschungs-Informationssystem (UFIS) als operationale Basis für Scaling-Konzepte in der Umweltplanung, in: Jaeschke, A. u.a. (Hrsg.), Informatik aktuell, Tagungsband für 7. Symposium Informatik für den Umweltschutz, Ulm, S. 427 - 440

Lenz, R./Herderich, C. (1993): Das Umweltforschungsinformationssystem UFIS - Ein integriertes Informationssystem über ökologische Vorhaben, Modelle und Daten, in: Denzer, Ralf u.a. (Hrsg.), 1. Workshop Integration von Umweltdaten. Bericht des Kernforschungszentrums Karlsruhe KfK Nr. 5187. Karlsruhe, S. 109-115

Springstubbe, Oliver (1993): Grundzüge der Klassifikation und systematische Beschreibung ökologischer Systeme als Gliederungsprinzip für ein Informationssystem ökologisch relevanter Daten und Modelle. Diplomarbeit am Lehrstuhl für Landschaftsökologie, Technische Universität München-Weihenstephan (unveröffentlicht). München, 146 S.

Weichselgartner, Erich (1993): World Wide Web, in: iX 12, Dezember 1993, S. 84-91

Weisemann, Ulrike (1994): Entwurf eines Datenbankschemas zur Dokumentation des Datenbedarfs und des Outputs von Simulationsmodellen in der Ökologie. Studienarbeit an der Abteilung Computergestützte Informationssysteme des Instituts für Angewandte Informatik der Technischen Universität Berlin (unveröffentlicht). Berlin, 61 S.